

**METHOD AND APPARATUS FOR PROVIDING MULTIPLE OUTPUT  
CHANNELS IN A MICROPHONE**

**BACKGROUND OF THE INVENTION**

**1. Field of the Invention**

5           The present invention relates to methods and apparatus for providing multiple output channels in a microphone. More particularly, the invention is concerned with the provision of an arrangement wherein a single microphone is adapted to produce one or more different audio outputs depending upon characteristics of a speaker or user of the microphone while facilitating a high degree of accuracy in the recognition of the user or  
15       speaker by the arrangement.

          Currently, in the technology wherein one or more speakers utilize a plurality of microphones at generally the same time, difficulties are encountered in being able to prioritize the particular microphone which is to be employed; in effect, actuated at any  
20       particular instance, or to be able to clearly distinguish or identify which speaker is utilizing any particular microphone at a specified point-in-time. Basically, the technology utilizes either an array of microphones which is designed to pick-up multiple speakers located within a predetermined confined space or room; for example, a conference room or auditorium, utilizing the microphone array in order to detect  
25       which particular speaker is most likely to be adapted to improve signal-to-noise ratio encountered within the specified room or confined space; or utilizing a microphone array in order to connect to a video system so as to track a speaker, especially during teleconferencing.

30       **2. Discussion of the Prior Art**

          Numerous patent publications are in existence which, in general, relate to the deployment of arrays of operatively associated microphones in order to be able to identify or recognize different speakers and/or prioritize the use of select microphones of the microphone arrays.

09927690-081001  
T00T80"0692260

Huang et al. U.S. Patent 6,173,059 B1 discloses a telephone system employing two or more microphones which are retained together and directed so as to face outwardly from a central point. Through the use of mixing circuitry, and controlled circuitry signals are combined and analyzed when received from the telephones, and the signal  
5 from one of the microphones, or from one or more predetermined combinations of microphone signals, are employed in order to track a speaker as the speaker moves about a room or various speakers situated about the room speak and then fall silent.

Anderson U.S. Patent 6,137,887 discloses a directional microphone system in which  
10 multiple microphone units are activated by a control system depending upon a speaker having his speech originate within a specified acceptance angle which is located in front of the microphones. This automatically identifies the microphone which provides for the best reception of the speaker, and in one instance only turns on one microphone for each speaker, and in other instances also allowing several microphones to turn on  
15 simultaneously for several talkers at predetermined points-in-time.

Martin et al. U.S. Patent 6,069,963 discloses a hearing aid having a multidirectional sensitivity based on the use of microphones positioned on the hearing aid, thereby enabling sounds to be received and determined at differences in sound transit time  
20 within a sound channel.

Nakazawa U.S. Patent 6,069,961 discloses a system utilizing multiple microphones which are adapted to detect the direction of a sound source and extracting therefrom an object sound with a high signal/noise ratio at an excellent degree of accuracy.  
25

Nagata U.S. Patent 6,009,396 discloses a method and system for microphone array input which provides for speech type recognition using band-pass power distribution for sound source position and direction estimation.

30 Baker U.S. Patent 5,686,957 pertains to a teleconferencing imaging system including automatic camera steering relative to the reception of sounds by a plurality of

microphones in an array connected to a voice-directional camera imaging system, the latter of which electronically selects segmented images from a selected panoramic video screen arranged around a conference table.

- 5 Bowen et al. U.S. Patent 5,625,697 discloses a microphone selection process for use in a multiple microphone voice actuating switching system, whereby, predicated on different qualities of speech signals as received in a plurality of microphones, this will enable the selection of the best received speech signals within the environment of a conference room.

10

Addeo et al. U.S. Patent 5,335,011 discloses a sound localization system for teleconferencing by employing self-steering microphone arrays, wherein a signal selection is implemented for the best video and sound image emanating from a virtual location on a displayed image.

15

Julstrom U.S. Patent 4,658,425 discloses a microphone actuating control system suitable for teleconference systems, wherein a selection is employed in conjunction with the different modulated signals indicating that an associated microphone of an array of microphone is the source of the first loudest microphone signal.

20

Finally, McDonnell et al. U.S. Patent 4,396,800 discloses a microphone switching device wherein a switch is positioned on a microphone handle so as to enable audio signals to be transferred by a user of the microphone from one location to a different location, particularly when the microphone is used on a soundstage or public address system. However, there is no disclosure of an encoding and decoding arrangement being incorporated into the microphone, as is the case of the present invention.

25

In the technology, none of these systems and arrangements of multiple phones, with the exception of the use of a switch to activate a signal as is disclosed in the microphone of McDonnell et al. U.S. Patent 4,396,800, provide for a single microphone enabling the utilization of multiple output channels for preferred utilized voice recognition.

30

## SUMMARY OF THE INVENTION

In essence, the present invention provides for a method and arrangement in creating a microphone adapted to produce one or more different audio streams or outputs depending upon the speaker presently using the microphone. In effect, this can be readily implemented by a main user or speaker, such as an interviewer on a radio or TV talk show, or any speaker in a conference room, intending to control the audio output streams by suitably activating a button or switch. This can be readily constituted of a mercury balance switch which is located in the microphone and is adapted to detect a microphone angle or orientation, or and alternatively, can be implemented by introducing or adding multiple microphone pick-up elements in the head of the microphone so as to enable energy/volume levels to be employed in order to detect the identity of the user or speaker.

Moreover, the microphone can be provided with a set of LEDs to provide visual feedback to the speakers indicating as to which particular channel is active. Also the output of any channel number of; for example 1 to N, can be encoded by utilizing multiple output wires, by adding a DC bias, or using modulation on different carrier frequencies.

In a physical application, it is possible to contemplate a speaker talking with or an interviewer interviewing another person, or persons, wherein the conversation is to be concurrently and practically instantaneously translated into a plurality of different languages, and then to have the resulting output audio in each language synchronized back to a video.

Consequently, it is imperative that high quality speech recognition be obtained as rapidly as possible. The speaker or interviewer, who is normally the primary user of the microphone, is ordinarily a good speaker who could be well trained in a speech recognition system, whereas in contrast therewith the person being addressed or interviewed (interviewee) will not be likely well trained, so one would require a more general statistical model for speech recognition. Moreover, the words and grammatical

usage of the interviewer and the interviewee (or interviewees) are likely to be quite different, and consequently it would be advantageous to provide a different speech recognizer for the interviewer or interviewee. Although there are basically two ways to implement the foregoing, such as in either hardware or software, primarily the  
5 technology has heretofore focused on software solutions to this problem, in an area of the technology currently referred to as "speaker identification".

In essence, "speaker identification" which is utilized in connection with software is subject to two problems. Firstly, the speaker identification introduces a time delay,  
10 whereby at any time the interviewee might wish to interject some comments and the interviewer would then "pass the microphone" to the interviewee. Consequently the speaker I.D. have to be continuously implemented, introducing a several second delay in time. Secondly, the speaker identification or I.D. is subject to mistakes, especially if the interview takes place in a noisy or poor sound transmissive environment.

15 To the contrary, in comparison with the use of software, employing a hardware solution is a much more rapid and reliable solution to the above-mentioned problems. There are two approaches, in which a first approach requires the interviewer to manually control the output of the microphone, either by pressing a button, switch or some other tactile  
20 device, or by adjusting the angle or orientation of the microphone to thereby automatically change the output. Another approach would be to install multiple pick-up elements in the head of the microphone, to additionally use energy pick-up elements in the head of the microphone, and to also use an energy-volume-direction information of an input signal in order to determine whether the speaker is or is not the person holding  
25 the microphone. A still further even more advanced solution could be employed in order to detect frequency vibrations produced in the hand of the user of the microphone during periods of speech indicating that the interviewer is the person speaking. Thereafter, the outputted microphone can be adjusted to identify the person speaking, and this can be implemented in a single channel by adding a DC bias or by modulating  
30 the signal on different carrier frequencies, or by using a pulsed signal to indicate that a

new speaker is talking. Furthermore, this may be also be implemented on multiple channels by the provision of more than one output wire.

Moreover, it is also possible to contemplate implementing an encoding by employing a pulsed signal instead of a DC bias, carrier frequency or two wires. Thus, in essence, rather than using a high or low frequency continually, whenever the microphone detects that someone else besides the user is speaking, this can place an invisible or inaudible "beep" on the line, which can be detected by the decoder, thereby saving battery life.

In essence, any acceptable stereo transmission technique in the art can be readily employing in connection with the foregoing.

In effect, the control of the microphone can be implemented by different methods, such as, through:

- 1) providing a tactile switch which is controlled by the interviewer or primary speaker, such as a button, trigger or toggle switch located on the microphone;
- 2) employing an angle sensor in connection with the microphone in order to detect the angular orientation thereof for selecting the voice modulated output;
- 3) utilizing a frequency detector whereby the interviewer is holding the microphone in order to recognize that it is the interviewer speaking by detecting vibrations in the hand holding the microphone;
- 4) locating multiple pick-up elements in the head of the microphone in order to detect as to whether the speech is emanating from the interviewer or the interviewee;
- 5) mounting an inexpensive camera on the microphone to be able to detect the lip motion of the user which can identify the speaker.

The microphone may be adapted to adjust the pick-up elements in any way which produces high-quality separation between the different speech patterns, and the



1 of the drawings representing a flowchart in a diagrammatic arrangement for providing multiple output channels in a single microphone.

#### **DETAILED DESCRIPTION OF A PREFERRED EMBODIMENT**

5 Referring to the flowchart 10 illustrated in the drawings, a microphone 12 is represented which receives an audio signal responsive to use thereof by a speaker. The microphone 12 is adapted to an apparatus 14 which determines the identity of the speaker utilizing the microphone, such as a speaker sensor 16, which components may be arranged within the confines of the actual microphone 12.

10

The microphone 12 may incorporate either a switch 20 which is in the form of a manual switch controlled by the speaker, or the current user of the microphone, or a position switch such as mercury switch which can determine the direction in which the microphone is facing during use thereof; or a sound or other electrical sensor or sensors which is or are arranged in a handle or gripping portion of the microphone, and which can be employed in order to detect when the current holder of the microphone is speaking in contrast with a non-holder of the microphone; or a clip fastened to a lapel on the clothing or located on the body of the speaker, and which is connected to the hand-held microphone through either a thin wire or in a wireless mode. This clip on the speaker may only be required to help detect the holder of the microphone as the person presently speaking, the audio of the small microphone is not used, whereas the hand-held microphone audio is that which is employed.

Upon the sensor 16 determining which of two or more speakers are utilizing the microphone 12, the audio signal 22 captured by the microphone 12 is encoded with a specified speaker indicator number 24 as determined by a speaker sensor in the encoder 26, which is also located in the microphone 12. The most common encoding would be either a high or low frequency bias, whereas another method which employable would be the use of a stereo wire (not shown) with two channels and to encode on different channels; also stereo encoding and possibly employing a pulse.



0927600-081001

The encoded signal is received by an audio card, whereupon the original audio signal is extracted and the speaker indicator number 24 decoded in a decoder 28. The speaker indicator number 24 is then available for the particular application which can make use of this in any manner as required, and pursuant to the invention can be employed for different speech recognition models so as to improve the accuracy of a well trained interviewer and of a speaker indicator interviewee.

The foregoing can be also employed in a microphone 12 which encodes the output audio signal 22 so as to provide two or more different channels to afford a choice as to which speech recognition model to employ by either a switch or toggle to select the channel; or a position switch installed in the microphone; or intensity of sound levels are measured via sensors located where the user is holding a microphone.

Installed in or attached to the microphone 12 can also be inexpensive camera 30. This camera is adapted to visually detect lip motion in order to identify the person who is speaking.

In an aspect where an additional clip on the microphone 12 may be positioned on one of the speakers and the output audio signal from the main microphone is encoded with a channel, in the event that the energy of the microphone on the speaker exceeds a threshold, then the encoding may be accomplished by adding a DC bias; or by adding a high frequency overtone; or may be by detecting the encoding in a speech recognizer and using a different speech recording model based on this encoding; where the encoding is recognized by a DC or low-frequency bandpass filter; or where the encoding is recognized by a high-frequency bandpass filter.

Alternatively, the encoding can be implemented by employing a pulsed signal instead of the DC bias, carrier frequency or two wires. Thus, in essence, rather than using a high or low frequency continually, whenever the microphone 12 detects that someone else besides the user is speaking, this can place an invisible or inaudible "beep" on the line, which can be detected by the decoder 28, thereby saving battery life. Hereby, any

acceptable stereo transmission technique known in the art can be readily employed in connection with the foregoing.

- 5 From the foregoing it becomes readily apparent that the invention clearly eliminates the need for the employing arrangements utilizing multiple microphones or complex software speaker identification modules and systems, and enables a particular multiple output channel to be provided in a single microphone in a simple and expedient manner at low cost and at a high efficiency in the operation thereof.
- 10 While the invention has been particularly shown and described with respect to preferred embodiments thereof, it will be understood by those skilled in the art that the foregoing and other changes in form and details may be made therein without departing from the spirit and scope of the invention.